Modeling the effect of 3 missense AGXT mutations on dimerization of the AGT enzyme in primary hyperoxaluria type I

UNIVERSITÀ DEGLI STUDI DI TORINO

Modeling the effect of 3 missense *AGXT* mutations on dimerization of the AGT enzyme in primary hyperoxaluria type I

Short Title: modeling AGXT mutations in PH1

Angela Robbiano 1§, Vladimir Frecer 2§, Jan Miertus 3, Cristina Zadro 4, Sheila Ulivi 4, Elena Bevilacqua 4, Giorgia Mandrile 1, Mario De Marchi 1, Stanislav Miertus 5, Antonio Amoroso 6*

1 Department of Clinical and Biological Sciences, University of Torino, Regione Gonzole, 10, 10043 Orbassano, Italy

2 Cancer Research Institute, Slovak Academy of Sciences, Vlárska, 7, SK-83391 Bratislava, Slovakia

3 Molecular Pathology, International Center for Genetic Engineering and Biotechnology, AREA Science Park, Padriciano, 99, I-34012 Trieste, Italy

4 Institute for Maternal and Child Health – IRCCS "Burlo Garofalo" and Medical Genetics Unit, University of Trieste, Via dell'Istria, 65/1, I-34127 Trieste, Italy

5 International Center for Science and High Technology UNIDO, AREA Science Park, Padriciano, 99 – 34012 Trieste, Italy

6 Department of Genetics, Biology and Biochemistry, University of Turin, Via Santena, 19, 10126 Turin, Italy

§ contributed equally to this work.

*** Corresponding author**

Professor Antonio Amoroso
Department of Genetics, Biology and Biochemistry
University of Turin
Via Santena 19
I-10126 Turin
Italy
Phone: +39-011-633.6760
Fax: +39-011-633.6521
e-mail: antonio.amoroso@unito.it

Disclosures:

Financial support: This study was partly supported by a 40% and 60% projects of MURST and by IRCCS Burlo Garofolo di Trieste "progetto ricerca corrente". The financial support of “Cassa di Risparmio di Torino” and Ricerca Sanitaria Finalizzata Regione Piemonte 2008 is also acknowledged.

Experimental investigation on human subjects: Informed consent for genetic testing was obtained from all patients following institutional rules and in adherence with the Declaration of Helsinki.

Conflict of interest: none

Abstract

Introduction

Mutations of the *AGXT* gene encoding the alanine:glyoxylate aminotransferase liver enzyme (AGT) cause primary hyperoxaluria type 1 (PH1). Here we report a molecular modeling study of selected missense *AGXT* mutations: the common Gly170Arg, the recently described Gly47Arg and Ser81Leu variants, predicted to be pathogenic using standard criteria.

Methods

Taking advantage of the refined 3D structure of AGT, we computed the dimerization energy of the wild type and mutated proteins.

Results

Molecular modeling predicted that Gly47Arg affects dimerization with a similar effect to that shown previously for Gly170Arg through classical biochemical approaches. In contrast, no effect on dimerization was predicted for Ser81Leu. Therefore, this probably demonstrates pathogenic properties via a different mechanism, similar to that described for the adjacent Gly82Glu mutation that affects pyridoxine binding.

Conclusion

This study shows that the molecular modeling approach can contribute to evaluating the pathogenicity of some missense variants that affect dimerization. However, *in silico* studies - aimed to assess the relationship between structural change and biological effects - requires the integrated use of more than one tool.

Key words:

AGXT gene, AGT enzyme, Primary Hyperoxaluria Type 1, molecular modeling

INTRODUCTION

Primary hyperoxaluria type I (PH1) is an autosomal recessive disorder caused by defects of the liver-specific peroxisomal enzyme alanine:glyoxylate aminotransferase (AGT, EC 2.6.1.44) encoded by the *AGXT* gene (1). The AGT molecule has a homodimeric structure of 2x43 kD/ 392 aminoacid subunits. AGT rapidly dimerizes after synthesis in the cytosol, and is imported into the peroxisomes, where it catalyzes the transamination of glyoxylate to glycine, coupled with the conversion of alanine to pyruvate, with pyridoxal-5-phosphate as a cofactor.

Deficiency of AGT in PH1 patients impairs the hepatic detoxification of glyoxylate and results in increased oxalate and glycolate in plasma and urine. Excess oxalate saturates body fluids and accumulates as insoluble calcium salts in several organs (2, 3). The natural history of untreated PH1 is characterized by progressive decline in renal function, from nephrolithiasis nephrocalcinosis to generalized oxalosis and death (4).

A comprehensive algorithm for the clinical and biochemical diagnosis of hyperoxaluria has been recently proposed (5). *AGXT* gene testing has become an important diagnostic tool (6-8) and provided that the pathogenicity of mutations is verified, it renders the invasive liver biopsy (for direct assaying of AGT activity) unnecessary in most cases (9). Conversely, when only private variants of unknown pathogenicity are found, molecular diagnosis remains doubtful.

The mutational spectrum of the *AGXT* gene encompasses almost 100 different mutations so far, including nonsense, splice site, frameshift and missense point mutations, and rarer entire exon deletions (6,8,10). Missense mutations may have various pathogenic effects. The most frequent, Gly170Arg, causes a remarkable trafficking defect, whereby the newly synthesized AGT subunit is diverted away from its normal peroxisomal location, in part degraded, while the remaining in part is diverted to the mitochondria where eventually dimerizes (11). This effect is synergistically

influenced by the common Pro11Leu polymorphism whose Leu allele introduces a mitochondrial signal sequence (12), that targets a fraction of the AGT protein to the mitochondrion (13), and decreases the enzyme activity by 30% (14). Docking of the correctly folded monomers into stable dimers represents a critical step of AGT biosynthesis. Patients carrying Gly170Arg, or other mistargeting mutations synergic with the Leu11, like Phe152Ile (14), are most likely to respond to vitamin B6 (Pyridoxine) (16-18). Other missense mutations are associated with accelerated degradation or intra-peroxisomal aggregation, that probably result from misfolding at later stages of AGT biosynthesis (14). Finally, specific mutations of critical residues in the catalytic- or the cofactor-binding sites can affect the enzymatic activity without altering AGT dimerization or targeting (19).

Determination of the crystal structure of the AGT enzyme has enabled us to improve the understanding of the pathogenic effects of some recurrent missense mutations, in terms of folding, dimerization and stability (20). Such studies may open the way to new therapeutic strategies, based on the structure-based design of small molecules capable of rescuing the defective enzyme, and thus extending the chance of conservative treatment to other more severe, life-threatening and non-Pyridoxine responsive mutations (21).

In this paper we approach the question by using a new *in silico* molecular modeling method aimed to rationalize the effects of specific missense AGT mutations on dimerization thermodynamics and provide proof of principle evidence of its performance.

SUBJECTS AND METHODS

Patients in whom the modeled mutations were identified belong to the Italian multicenter study of primary Hyperoxaluria, and had given their informed consent to use their genetic and clinical data for research purposes, as already described (17,22). Age at onset, relevant clinical data and

residual liver AGT activity, when available, are shown in Table 1. Screening of *AGXT* gene mutations was performed on genomic DNA through PCR amplification and direct sequencing, whereas healthy controls were analysed by DHPLC. The refseq NP_000021.1 protein sequence of the human AGT gene was used as reference. Multiple sequence alignment of 22 orthologous species was generated using ClustalW (http://www.ebi.ac.uk/Tools/clustalw2/index.html, Supplementary material, Fig.2). The evolutionary conservation at the mutated positions was evaluated using the SIFT (23), SNPs3D (24) and PMUT (25) algorithms available on line. The "major" or "minor" haplotype were characterized by inspection of the relevant sequence features at codons 11 and 340 (26).

Molecular Modeling. Starting from the crystal structure of the *wt* AGT homodimer ***WT:WT*** (20) we prepared the refined three-dimensional (3D) structure of the wild type (*wt*) dimer bound to the pyridoxal-5'-phosphate (PLP) cofactor and the aminooxyacetic acid (AOA) inhibitor (Figure 1) available from *Protein Data Bank* (27) - and from there we prepared the structure of the *wt* AGT monomer.

In the 3D models of the mutant AGT forms (*mt*) we replaced the side chains at the mutated sites by the best rotamer of the substituting residue followed by careful conformational search and geometry optimization using the molecular modeling software *Insight-II* (Accelrys Software Inc., San Diego, CA, release 2000). We computed the conformational and interaction energies of each AGT monomer and dimer by molecular mechanics (MM) using class II consistent force field CFF91 and charge parameters (28).

The dimerization of two *mt* AGT monomers in aqueous solution to form a homodimer (***MT:MT***) can be represented by the energy of dimerization: $\Delta E_{tot} = E_{tot}[MT{:}MT] - 2E_{tot}[MT]$, composed of contributions from molecular mechanics (MM) potential energy (ΔE_{MM}) and solvent effects

(ΔE_{solv}) as: $\Delta E_{tot} = \Delta E_{MM} + \Delta E_{solv}$. When computing dimerization energy, we took into account the interaction between monomers in the dimer, the stability and molecular structure of the free monomers and the effect of solvent upon monomers association. We compared the different *mt* AGT forms via relative changes in the dimerization energy, $\Delta\Delta E_{tot} = \Delta E_{tot}[MT{:}MT] - \Delta E_{tot}[WT{:}WT]$ with respect to the reference *wt* homodimer *WT:WT*. Relative changes were defined in a similar way for the contributions of both the MM interaction energy and the solvent effect.

When computing the solvation energy, we also incorporated the effects of ionic strength through the solution of nonlinear Poisson-Boltzmann equation (29), using the software package *DelPhi* (Accelrys Software Inc., San Diego, CA, release 2000).

We calculated the molecular solvent-accessible surface area of the *wt* and *mt* residues by using the Connolly algorithm (30) implemented in the *Insight-II* software. Relative changes in the accessible surface area of the studied residues upon dimer formation ($\Delta\Delta S_{bur}$) were calculated as a difference in the solvent accessible surfaces of residues in the dimer and in the free monomers with respect to the corresponding residues in the reference *wt* form of AGT: $\Delta\Delta S_{bur} = (2S_{Con}[mt]_{MT} - S_{Con}[mt,mt]_{MT:MT}) - (2S_{Con}[wt]_{WT} - S_{Con}[wt,wt]_{WT:WT})$.

RESULTS

AGXT Mutations. The 3 selected mutations were all absent in 80 healthy ethnically-matched controls (160 chromosomes). Multiple sequence alignment showed evolutionary conservation at residue Gly47 and Gly170 in 22/22 species and at Ser81 in 17/22 (Supplementary material, Fig.2). The SIFT and SNPs3D software tools concordantly predicted Gly47Arg and Ser81Leu to be not tolerated, with similar or stronger scores than the common Gly170Arg mutation (Table 1),

and were considered as pathogenic. In contrast, the PMUT software predicted only Gly47Arg to be pathogenic, and Ser81Leu to be tolerated.

We then implemented a molecular modeling approach to study the effect of the three missense changes on AGT dimerization, using the known Gly170Arg mutation as a positive control. For this, we modeled the homodimeric AGT forms of each of the three changes, as well as the heterodimer expected in the Gly170Arg/Gly47Arg compound heterozygous patient 1. No heterodimer was expected in patient 2 who carried a null mutation on the second allele.

Molecular Modeling. Residue Gly47 is located on the dimerization interface of AGT, close to the C-terminal boundary of the long, irregular N-terminal coil (residues 1-21) that upon dimerization wraps around the cognate monomer (Figure 2). Residue Ser81 is similarly positioned on the dimerization interface, near the cofactor binding site. In contrast, residue Gly170 is located on the exposed surface without direct contact with the cognate monomer. Due to its position at the dimer interface, the substitution of the small Gly47 residue by a bulky cationic Arginine contributes more to the relative surface area buried upon AGT dimerization ($\Delta\Delta S_{bur}$) than the other mutations that either insert a smaller residue or are located outside of the dimerization interface (Table 2). The Gly47Arg and Gly170Arg substitutions stabilize the AGT homodimers via increased monomer-monomer interaction, $\Delta\Delta E_{MM}$, more than the Ser81Leu substitution (Table 2).

On the other hand, the attractive MM interactions caused by the non conservative Gly to Arg substitutions are largely compensated for upon AGT dimerization by the effect of the solvent. The *wt* AGT monomer bears a total molecular charge of $+1e^-$, while Arg47 and Arg170 increase its charge to $+2\ e^-$. In both cases, and especially for Arg47, the free *mt* monomers with the charged Arg residues exposed to the solvent are better stabilized by the interactions with water

($\Delta\Delta E_{solv}$) than the dimers, where the corresponding residues are partially buried (Table 2). Thus the stability of the mutated dimers is significantly diminished by the effect of protein hydratation. In contrast, the Ser81Leu substitution does not change the charge. The overall dimer stability, which takes into account both the interactions between the residues of the associated monomers and the solvent effect ($\Delta\Delta E_{tot}$), is weaker for all the considered *mt* dimers than the native ***WT:WT***. Our computational approach predicts a stronger decrease of dimer stability for the *mt* ***G47R:G47R*** *>* ***G170R:G170R*** dimers compared to the ***WT:WT*** reference, while only a minor change in dimer stability is predicted for the *mt* ***S81L:S81L*** homodimer (Table 2).

DISCUSSION

The pathogenic effect of recurrent AGT mutations has been extensively investigated with classical biochemical and biological approaches. A series of studies, both *ex vivo* and *in vitro*, have highlighted the peculiar effect of the recurrent Gly170Arg mutation on AGT biosynthesis (15). However, no comparable knowledge is available for the remaining missense AGT mutations, about 30-34% in Caucasians, in which a growing series of rarer or private changes are identified (10). In these cases, the current criteria of cosegregation of disease in families and lack of disease in controls does not always allow to discriminate true mutations from rare polymorphisms. The liver specific expression of AGT further hampers to evaluate the effect of promoter and splice-control mutations on mRNA biosynthesis. For missense mutations, determination of the crystal structure of AGT has opened the way to rationalize the pathogenic mechanisms in terms of folding, dimerization and stability of the enzyme (20). Here we describe the application of molecular modeling techniques as a means to explore the pathogenic potential of AGT missense mutations. Similar approaches was previously employed by Burnett et al (31)

for hemoglobin and Kobayashi et al (32) for lipoprotein lipases, but to our knowledge their use for AGT mutations is novel.

We assumed the working hypothesis that some mutations, in particular the Gly47Arg situated on the dimerization interface, affect the formation of AGT dimers and may exert effects similar to those demonstrated for Gly170Arg. Indeed, when we rationalized the stability of the mutated Gly170Arg and Gly47Arg dimers in terms of net charge and size of the replacing Arg residue and its location on the molecular surface with respect to the dimerization region, solvation resulted to exert a dominant effect and stabilize the free monomers more than the corresponding dimers. In our molecular modeling both mutations change the $\Delta\Delta E_{tot}$ by an order of tenth kcal·mol^{-1} , i.e. shift the equilibrium constant of the *mt* molecules from dimer to free monomeric *mt*. In other words, dimers with either mutation are predicted to be less thermodynamically stable at physiological conditions than the wild type AGT dimer.

Decreased stability of the mutated AGT dimers has been previously considered as a possible cause of the peroxisomes-to-mitochondria mistargeting (15) and accelerated AGT degradation or aggregation that have been associated with the Gly170Arg mutation (14). According to our modeling study, Gly47Arg destabilizes the AGT dimer even more than Gly170Arg. Interestingly, a similar destabilizing effect is predicted for both the Gly170Arg:Gly170Arg and Gly47Arg:Gly47Arg mutated homodimers and for the Gly47Arg:Gly170Arg heterodimer, whose corresponding *in vivo* situation is represented by the compound heterozygous patient 1. On overall, these data support the pathogenicity of Gly47Arg, although its clinical severity cannot be established in the presence of the mild Gly170Arg mutation, as in general true for new described mutations when found in compound heterozygosity with mild ones.

The effect of the Ser81Leu mutation predicted from molecular modeling is much less pronounced, with a minimal change of the $\Delta\Delta E_{tot}$, suggesting that secondary structures and dimers formation are not affected. This is in agreement with the PMUT results, based also on the profile of secondary structure and solvent accessibility. On the other hand, Ser81Leu should be considered pathogenic on the basis of absence in ethnically matched controls and high degree of evolutionary conservation. Its proximity to the cofactor and substrate binding sites suggests that it may affect the catalytic function of the enzyme. Indeed, the adjacent Gly82Glu mutation was previously described to affect pyridoxine binding (14).

The wide mutational spectrum of primary hyperoxaluria and the consequent variety of the AGT enzyme defects makes the use of multiple approaches to characterize novel mutations highly recommended. Since dimerization represents a critical step in AGT biosynthesis, the presented approach can significantly contribute to assess the pathogenicity of missense mutations, clarifying their contribution to the phenotype and providing some insight into the structural and functional changes of AGT responsible of the PH1 phenotype.

Acknowledgement

Nephrologists who referred patients are gratefully acknowledged: Licia Peruzzi (OIRM-S.Anna Hospital, Torino, Italy), Martino Marangella (Mauriziano Hospital, Torino, Italy).

REFERENCES

1. Purdue PE, Lumb MJ, Fox M, et al. Characterization and chromosomal mapping of a genomic clone encoding human alanine:glyoxylate aminotransferase. Genomics. 1991; 10: 34-42.
2. Marangella M, Petrarulo M, Vitale C, et al. The primary hyperoxalurias. Contrib Nephrol. 2001; 136: 11-32.
3. Danpure CJ, Jennings PR. Peroxisomal alanine:glyoxylate aminotransferase deficiency in primary hyperoxaluria type I. FEBS Lett. 1986; 201: 20-24.
4. Milliner DS, Wilson DM, Smith LH. Clinical expression and long-term outcomes of primary hyperoxaluria types 1 and 2. J Nephrol. 1998;11 Suppl 1: 56-59.
5. Milliner DS. The primary hyperoxalurias: an algorithm for diagnosis. Am J Nephrol. 2005; 25: 154-160.
6. Williams E, Rumsby G. Selected exonic sequencing of the AGXT gene provides a genetic diagnosis in 50% of patients with primary hyperoxaluria type 1. Clin Chem. 2007; 53: 1216-1221.
7. Rumsby G, Williams E, Coulter-Mackie M. Evaluation of mutation screening as a first line test for the diagnosis of the primary hyperoxalurias. Kidney Int. 2004; 66: 959-963.
8. Monico CG, Rossetti S, Schwanz HA, et al. Comprehensive mutation screening in 55 probands with type 1 primary hyperoxaluria shows feasibility of a gene-based diagnosis. J Am Soc Nephrol. 2007; 18: 1905-1914.

9. Danpure CJ, Jennings PR, Watts RW. Enzymological diagnosis of primary hyperoxaluria type 1 by measurement of hepatic alanine: glyoxylate aminotransferase activity. Lancet. 1987; 1: 289-291.

10. Williams EL, Acquaviva C, Amoroso A et al. Primary hyperoxaluria type 1: update and additional mutation analysis of the AGXT gene. Hum Mutat. 2009; 30: 910-917.

11. Danpure CJ, Cooper PJ, Wise PJ, Jennings PR. An enzyme trafficking defect in two patients with primary hyperoxaluria type 1: peroxisomal alanine/glyoxylate aminotransferase rerouted to mitochondria. J Cell Biol. 1989;108: 1345-1352.

12. Purdue PE, Allsop J, Isaya G, Rosenberg LE, Danpure CJ. Mistargeting of peroxisomal L-alanine:glyoxylate aminotransferase to mitochondria in primary hyperoxaluria patients depends upon activation of a cryptic mitochondrial targeting sequence by a point mutation. Proc Natl Acad Sci U S A. 1991; 88: 10900-10904.

13. Purdue PE, Takada Y, Danpure CJ. Identification of mutations associated with peroxisome-to-mitochondrion mistargeting of alanine/glyoxylate aminotransferase in primary hyperoxaluria type 1. J Cell Biol. 1990; 111: 2341-2351.

14. Lumb MJ, Danpure CJ. Functional synergism between the most common polymorphism in human alanine:glyoxylate aminotransferase and four of the most common disease-causing mutations. J Biol Chem. 2000; 275: 36415-36422.

15. Leiper JM, Oatey PB, Danpure CJ. Inhibition of alanine:glyoxylate aminotransferase 1 dimerization is a prerequisite for its peroxisome-to-

mitochondrion mistargeting in primary hyperoxaluria type 1. J Cell Biol. 1996; 135: 939-951.

16. van Woerden CS, Groothoff JW, Wijburg FA, Annink C, Wanders RJ, Waterham HR. Clinical implications of mutation analysis in primary hyperoxaluria type 1. Kidney Int. 2004; 66: 746-752.

17. Amoroso A, Pirulli D, Florian F, et al. AGXT gene mutations and their influence on clinical heterogeneity of type 1 primary hyperoxaluria. J Am Soc Nephrol. 2001; 12: 2072-2079.

18. Monico CG, Olson JB, Milliner DS. Implications of genotype and enzyme phenotype in pyridoxine response of patients with type I primary hyperoxaluria. Am J Nephrol. 2005; 25: 183-188.

19. Purdue PE, Lumb MJ, Allsop J, Minatogawa Y, Danpure CJ. A glycine-to-glutamate substitution abolishes alanine:glyoxylate aminotransferase catalytic activity in a subset of patients with primary hyperoxaluria type 1. Genomics. 1992; 13: 215-218.

20. Zhang X, Roe SM, Hou Y, et al. Crystal structure of alanine:glyoxylate aminotransferase and the relationship between genotype and enzymatic phenotype in primary hyperoxaluria type 1. J Mol Biol. 2003; 331: 643-652.

21. Danpure CJ. Primary hyperoxaluria: from gene defects to designer drugs? Nephrol Dial Transplant. 2005; 20: 1525-1529.

22. Pirulli D, Puzzer D, Ferri L, et al. Molecular analysis of hyperoxaluria type 1 in Italian patients reveals eight new mutations in the alanine: glyoxylate aminotransferase gene. Hum Genet. 1999; 104: 523-525.

23. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001; 11: 863-874.

24. Yue P, Moult J. Identification and analysis of deleterious human SNPs. J Mol Biol. 2006; 356: 1263-1274.

25. Ferrer-Costa C, Gelpi JL, Zamakola L, Parraga I, de la Cruz X, Orozco M. PMUT: a web-based tool for the annotation of pathological mutations on proteins. Bioinformatics. 2005; 21: 3176-3178.

26. Pirulli D, Marangella M, Amoroso A. Primary hyperoxaluria: genotype-phenotype correlation. J Nephrol. 2003; 16: 297-309.

27. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. Nucleic Acids Res. 2000; 28: 235-242.

28. Maple J, Hwang M, Stockfish T, et al. Derivation of Class II Force Fields. 1. Methodology and Quantum Force Field for the Alkyl Functional Group and Alkane Molecules. J Comput Chem. 1994; 15: 162-182.

29. Gilson MK, Honig B. The inclusion of electrostatic hydration energies in molecular mechanics calculations. J Comput Aided Mol Des. 1991; 5: 5-20.

30. Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. Science. 1983; 221: 709-713.

31. Burnett JC, Botti P, Abraham DJ, Kellogg GE. Computationally accessible method for estimating free energy changes resulting from site-specific mutations of biomolecules: systematic model building and structural/hydropathic analysis of deoxy and oxy hemoglobins. Proteins. 2001; 42: 355-377.

32. Kobayashi Y, Nakajima T, Inoue I. Molecular modeling of the dimeric structure of human lipoprotein lipase and functional studies of the carboxyl-terminal domain. Eur J Biochem. 2002; 269: 4701-4710.

Legends for figures

Figure 1

Dimerization site/active site view of the relaxed *wt* AGT monomer in a ribbon rendering. The co-factor pyridoxal-5'-phosphate (PLP) and the substrate aminooxyacetic acid (AOA) are shown in Corey-Pauling-Kultun (CPK) representation in cyan and brown. The locations of the Leu11 and Met340 residues of the minor polymorphic form are indicated by horizontal and vertical arrows respectively. The mutated residues are highlighted: Gly47 in yellow, Ser81 in green and Gly170 in white.

Figure 2

(A) Front view and **(B)** side view of the AGT dimer ***WT:WT*** in Corey-Pauling-Kultun (CPK) representation, with monomer A shown in red and monomer B in blue. Positions of the mutated residues GlyA47 and GlyB47 are shown in yellow, residues GlyA170 and GlyB170 in white. Residues SerA81 and SerB81 located at the dimerization interface are not visible. The long N-terminal extensions (coils of residues 1 – 21) wrap around the N-terminal segment of the opposing monomer.

TABLES

Table 1. Clinical, biochemical and molecular findings of PH1 patients.

Patient ID gender (origin)	**Age at onset (yrs)**	**AGT enzyme activity**	**cDNA**	**Protein**	***AGXT* haplotype***	***Evolutionary conservation***	**Pathogenicity score[§] SIFT - SNPs3D - PMut**		
1 male (Sicily)	35	2%	c.139G>A	G47R	minor	22/22	0.00 non toler.	-1.91 non toler	0.59 non toler
			c.508G>A	G170R	minor	22/22	0.00 non toler.	-0.55 non toler	0.56 non toler
2 female (Serbia)	1	Not done	c.242C>T	S81L	maior	17/22	0.01 non toler	-0.51 non toler	0.31 tolerated
			c.614C>A	S205X	maior				

* minor haplotype is defined by P11L + 74bp-duplication in intron 1 + I340M.

§ non tolerated: SIFT < 0.05 (23), SNPs3D < 0 (24) PMut > 0.5 (25).

Table 2. Buried surface and relative energies of dimerization of wild type (*wt*) and mutated (*mt*) human AGT monomers.

AGT dimer	**Interacting monomers**		**Relative buried surface [Å^2]**	**Relative energy of dimerization [kcal·mol^{-1}]**		
	Monomer MT_i	**Monomer MT_j**	$\Delta\Delta S_{bur}$ [a]	$\Delta\Delta E_{MM}$ [b]	$\Delta\Delta E_{solv}$ [c]	$\Delta\Delta E_{tot}$ [d,e]
wt:wt *WT:WT* homodimer	*wt* AGT (Q_{wt} = +1) [f]	*wt* AGT (Q_{wt} = +1)	0	0	0	0
G47R:G47R $MT_i:MT_j$ homodimer	*mt* AGT *Gly47Arg* (Q_{mt} = +2)	*mt* AGT *Gly47Arg* (Q_{mt} = +2)	158	-23	106	83
G170R:G170R $MT_i:MT_j$ homodimer	*mt* AGT *Gly170Arg* (Q_{mt} = +2)	*mt* AGT *Gly170Arg* (Q_{mt} = +2)	44	-16	56	40
S81L:S81L $MT_i:MT_j$ homodimer	*mt* AGT *Ser81Leu* (Q_{mt} = +1)	*mt* AGT *Ser81Leu* (Q_{mt} = +1)	14	1	2	3
G47R:G170R $MT_i:MT_j$ heterodimer	*mt* AGT *Gly47Arg* (Q_{mt} = +2)	*mt* AGT *Gly170Arg* (Q_{mt} = +2)	110	-19	74	55

[a] $\Delta\Delta S_{bur}$ – relative contribution of mutated residue to the buried surface upon dimer formation was calculated as the difference in Connolly surfaces (30) of the

concerned residues in the dimer *MT:MT* and in free monomer *MT* with respect to the native residues in the homodimer *WT:WT* $\Delta\Delta S_{bur} = (2S_{Con}[mt]_{MT} - S_{Con}[mt,mt]_{MT:MT}) - (2S_{Con}[wt]_{WT} - S_{Con}[wt,wt]_{WT:WT})$, in [Å^2];

[b] $\Delta\Delta E_{MM} = (E_{MM}[MT:MT] - 2E_{MM}[MT]) - (E_{MM}[WT:WT] - 2E_{MM}[WT])$ is the relative molecular mechanics interaction energy contribution to the *MT:MT* AGT dimer formation with respect to the *wt* homodimer *WT:WT*;

[c] $\Delta\Delta E_{solv} = (E_{solv}[MT:MT] - 2E_{solv}[MT]) - (E_{solv}[WT:WT] - 2E_{solv}[WT])$ is the relative solvation energy contribution to the *MT:MT* AGT dimer formation with respect to the *wt* homodimer *WT:WT*;

[d] $\Delta\Delta E_{tot} = \Delta\Delta E_{MM} + \Delta\Delta E_{solv}$ is the relative total energy change connected with the *MT:MT* AGT dimer formation with respect to the *wt* homodimer *WT:WT*;

[e] a change of $\Delta\Delta E_{tot}$ by 10 kcal·mol^{-1} corresponds to about 3.2% change in the energy of dimerization the *WT:WT* dimer. For comparison, an increase of Gibbs free of a dimerization reaction by 10 kcal·mol^{-1} will cause a shift of the equilibrium constant of this reaction by a factor of $5 \cdot 10^8$ towards dissociated monomers;

[f] Q_{wt}, Q_{mt} are the molecular charges of *wt* or *mt* AGT monomers including the co-factor PLP and substrate AOA.

Figure 1

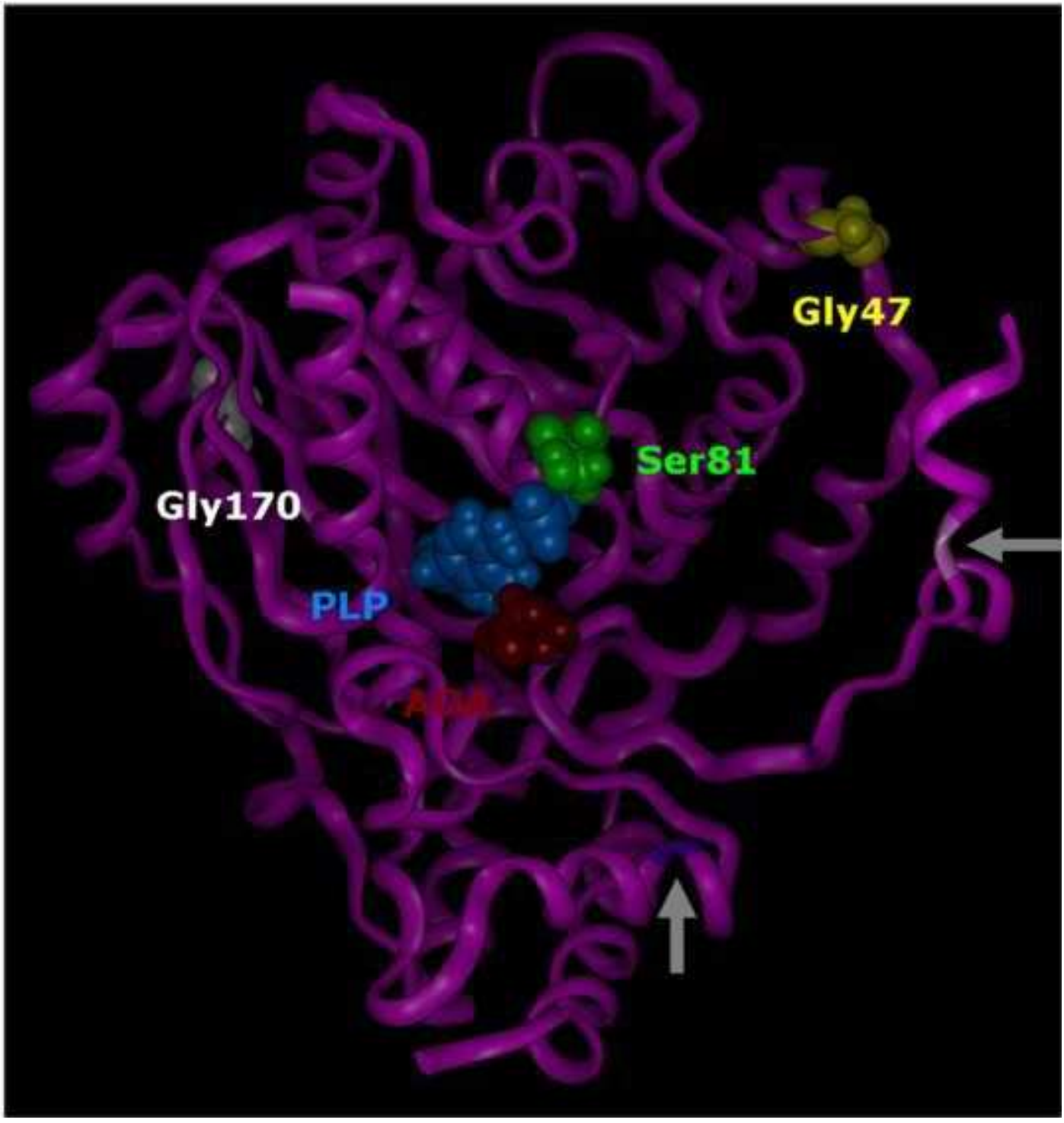
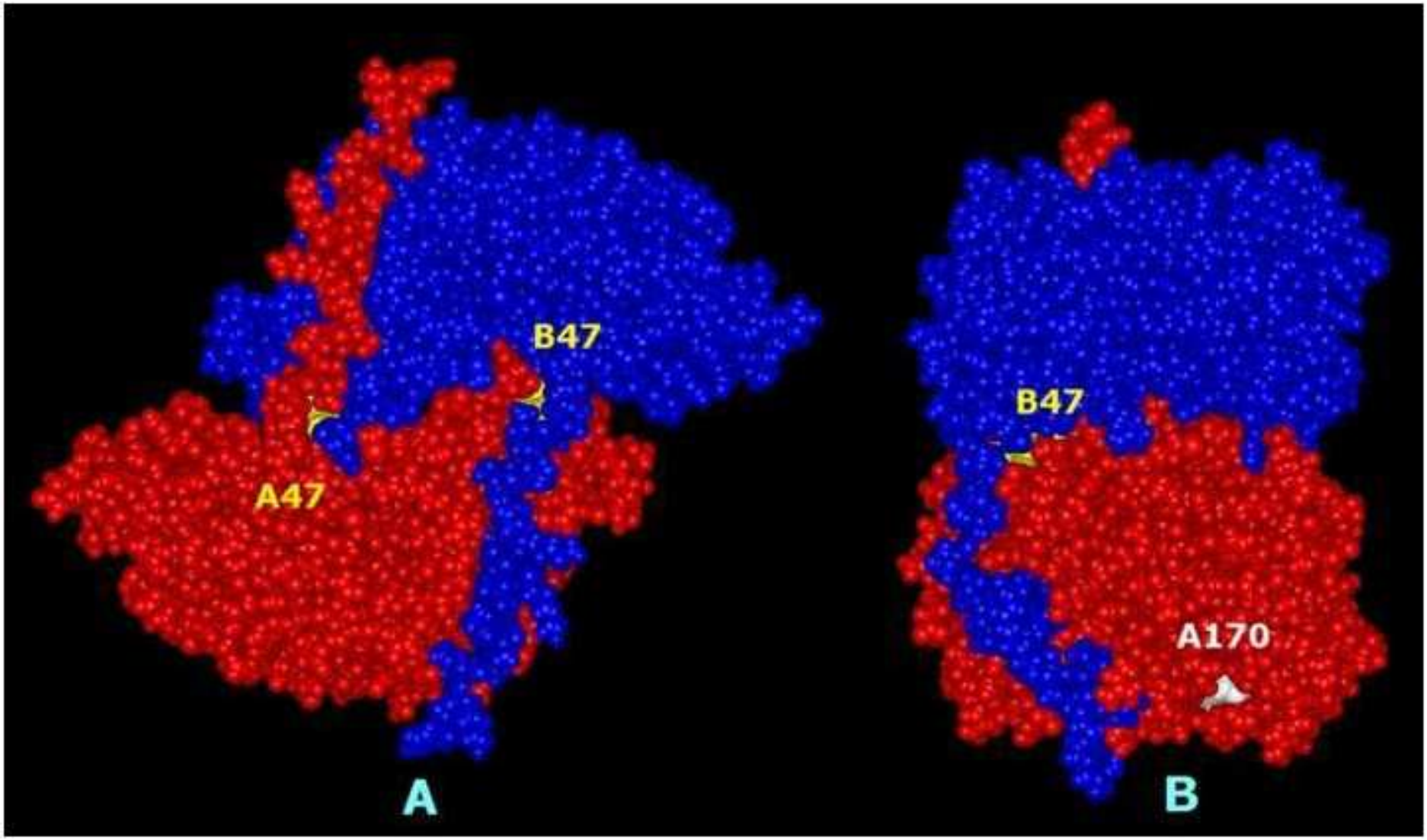


Figure2



Supplementary Material

Supplementary Material 1: Legend to Figure 1

Electropherograms of the three missense mutations studied

Supplementary Material 2.

Multiple sequence alignment of AGT orthologous sequences of 22 different species.

The four mutated residues are highlighted with boxes, and the substituting residue is indicated above the top line. The degree of evolutionary conservation of each residue is indicated below the bottom line, according to Clustal rules: * identity, : conservative substitution, . semi-conservative substitution

H.sapiens NP_000021.1, M.mulatta XP_001090301.1, P.pygmaeus Q5RDP0.1, C.jacchus P31029.1, C.familiaris XP_855949.1, F.catus NP_001036031.1, E.caballus XP_001497552.1, B.taurus NP_001095825.1, O.cuniculus NP_001075778.1, O.anatinus XP_001513675.1, R.norvegicus NP_085914.1, M.musculus NP_057911.2, X.laevis NP_001081948.1, D.rerio NP_001002331.1, D.melanogaster NP_511062.1, A.aegypti XP_001660875.1, N.vitripennis XP_001608017.1, A.mellifera XP_397119.2, C.elegans NP_495885.1, S.purpuratus XP_780986.2, Cyanothece sp. CCY0110 ZP_01727410.1, Prochlorococcus marinus str. NATL2A YP_291337.1

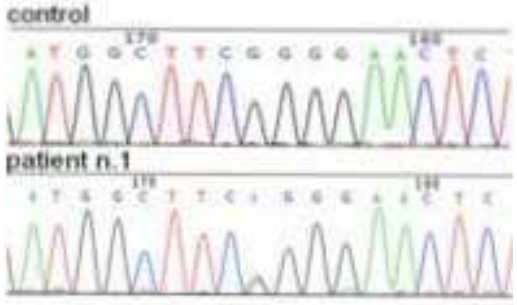
Multiple sequence alignment of AGT orthologs using CLUSTAL 2.0.2.

```
H.sapiens             ------------------------------------------------------------
M.mulatta             ------------------------------------------------------------
Pongo_pygmaeus        ------------------------------------------------------------
C.jacchus             ---------------------------------------------MFQALAKASAALGP 14
Canis_familiaris      ---------------------------------------------MFRALARASVALGP 14
Felis_catus           ---------------------------------------------MFRALARASATLGP 14
Equus_caballus        MPSQEAVADPGPPLLPCLPAACKGFCFLLTLQPLGLVWSRRGASREMLRALAMASVALGP 60
Bos_taurus            ---------------------------------------------MLWALTTARAVLGH 14
O.cuniculus           ------------------------------------------------------------
O.anatinus            ------------------------------------------------------------
Rattus_norvegicus     ------------------------------------------------------------
Mus_musculus          ---------------------------------------------MFRMLAKASVTLGS 14
Xenopus_laevis        ---------------------------------------------MQGSVRSISSSLLCA 15
Danio_rerio           ------------------------------------------------------------
D.melanogaster        ------------------------------------------------------------
Aedes_aegypti         ------------------------------------------------------------
N.vitripennis         ------------------------------------------------------------
Apis_mellifera        ------------------------------------------------------------
C.elegans             -----------------------------------------------------MISTRFLRP 9
```

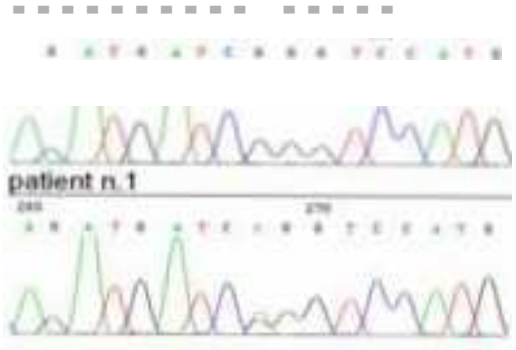
```
S.purpuratus          ------------------------------------------------------------
Cyanothece_bacter     ------------------------------------------------------------
Pr.coccus_marinus     ----------------------------------------MQSLKKTPRLCLSPFRLGRY 20

                                        L                                     R
H.sapiens             --------MASHKLLVTPPKALLKPLSIPNQLLLGPGPSNLPPRIMAAGGLQMIGSMSKD 52
M.mulatta             --------MASHKLLVPPPKALLKPLSIPKRLLLGPGPSNLPPRIMSAGGMQIIGPMDKE 52
Pongo_pygmaeus        --------MASHKLLVHPPKALLKPLSIPNRLLLGPGPSNLAPRIMAAGGLQIIGPMSKD 52
C.jacchus             RAAGWVRTMASHQLLVAPPKALLKPLSIPTRLLLGPGPSNLPPRTMAAGGLQMLGHMHKE 74
Canis_familiaris      QAAGWVRTMASQQLLVAPPKALLKPLSIPNRLLLGPGPSNLAPRILAAGGLQMIGHMHKE 74
Felis_catus           QVAGWARTMATCQLLVAPPEALLRPLSIPNRLLLGPGPSNLAPRVLVAGGKQMIGHMHKE 74
Equus_caballus        QAAGWARTMASHQLLVAPPEALRKPLSIPNRLLLGPGPANLAPRVLAAGGLQVIGHMHKE 120
Bos_taurus            RTAGWVRTMASHSLLVAPPAALSKPLFIPSRLLLGPGPSNLTPRVMAAGGLQVLGHMHQE 74
O.cuniculus           --------MASRQLLVAPPEALRKPLCTPHRLLLGPGPSNLPPRVLAAGGLQMIGHMHEE 52
O.anatinus            --------MSSQKLLVSPPPSLLKPFVIPDKLLLGPGPSNLPPRVMAAGGLQMIGHMHKE 52
Rattus_norvegicus     --------MGSHQLLVPPPEALSKPLSIPKRLLLGPGPSNLAPRVLAAGSLRMIGHMQKE 52
Mus_musculus          RAAGWVRTMGSYQLLVPPPEALSKPLSVPTRLLLGPGPSNLAPRVLAAGSLRMIGHMQKE 74
Xenopus_laevis        ARLSAPVRIMSSLATIPPPSALQRPLNVPQRLMLGPGPSNVPPRIQAAGGLQLIGHMHPE 75
Danio_rerio           ---------MSSLSVL-PPECLLKPFPVPQRLMLGPGPSNVPARISAAGAQPMLGHLHTE 50
D.melanogaster        -------------MEVPPPLVLKRPLYVPSKTLMGPGPSNCSHRVLEAMSNPVLGHMHPE 47
Aedes_aegypti         -----------MEYKVTPPAVLREPLVTPNKLLMGPGPSNAPQRVLDAMSRPILGHLHPE 49
N.vitripennis         -------------MEVEPPKELLKPLQLPQRILTGPGPSNCSQRVLKALEQQVLGHTHPE 47
Apis_mellifera        ----MNYKWNKLAVHKEPPQILRTKLQLPIKTLMSPGPTNCSKRVLQSLQNQILGHLHPE 56
C.elegans             SVSIFGFGIKSSMSSRAPPKALLQDMVVPPRQLFGPGPSNMADSIAETQSRNLLGHLHPE 69
S.purpuratus          -------MVGNKSMAVSSPKELLLPMIVPQKLLMGPGPSNVPPRILAAGALPILGHMHPE 53
Cyanothece_bacter     ---------MTIPSQP-------QPLTIPPRLLMGPGPSNANPRILSAMSLPAIGHLDPF 44
Pr.coccus_marinus     INRSVFLATQILPSVDNQHRKDFGPTVSPERLLLGPGPSNADPAVLKALSQPPIGHLDPF 80
                                          * : : .***:*       :     :*
                                          L
H.sapiens             MYQIMDEIKEGIQYVFQTRNPLTLVISGSGHCALEAALVNVLEPGDSFLVGANGIWGQRA 112
M.mulatta             MYQIMDEIKEGIQYVFQTRNPLTLVITGSGHCALEAALVNVLEPGDSFLVGVNGIWGQRA 112
Pongo_pygmaeus        MYQIMDEIKEGIRYVFQTRNPLTLVISGSAHCALEAALVNVLEPGDSFLVGANGIWGQRA 112
C.jacchus             TYQIMDEIKEGIQYVFQTRNPLTLVISGSGHCALEAALINVLEPGDSFLVGVNGIWGQRA 134
Canis_familiaris      MFQIMDEIKEGIQYVFQTKNPLTLAVSGSGHCALEAALFNLLEPGDSFLVGANGIWGQRA 134
Felis_catus           MFQIMDDIKQGIQYVFQTKNPLTLAISGSGHCALEAALFNILEPGDPFLVGVNGIWGQRA 134
Equus_caballus        MYQIMDEIKQGIQYVFQTRNPLTLAISGSGHCALEAALFNLLEPGDSFLVGVNGIWGQRA 180
Bos_taurus            VYQIMDEIKQGIQYVFQTRNPLSLAISGSGHCALEAALFNLLEPGDSFLVGVSGIWGQRA 134
O.cuniculus           MYQVMDEIKQGIQYAFQTRNALTLAVSGSGHCALETALFNLLEPGDAFLVGANGIWGQRA 112
O.anatinus            MFQIMDEIKEGIQYAFQTKNQITLAISGSGHSAMEAALFNALEPGDSVMVGVNGIWGQRA 112
Rattus_norvegicus     MFQIMDEIKQGIQYVFQTRNPLTLVVSGSGHCAMETALFNLLEPGDSFLVGTNGIWGIRA 112
Mus_musculus          MLQIMEEIKQGIQYVFQTRNPLTLVVSGSGHCAMETALFNLLEPGDSFLTGTNGIWGMRA 134
Xenopus_laevis        MFQIMDDIKQGIQYAFQTKNNLTFAVSGSGHCAMETAIFNVVEKGDVVLVAVKGIWGERA 135
Danio_rerio           TIEIMNQIKSGIQYAFQTKNRVTLAVSGPGHAAMECAIFNSLEPGDKILIAVNGIWGERA 110
D.melanogaster        CLQIMDEVKEGIKYIFQTLNDATMCISGAGHSGMEAALCNLIEDGDVVLMGITGVWGHRA 107
Aedes_aegypti         TLKIMDDIKEGVRYLFQTNNIATFCLSASGHGGMEATLCNLLEDGDVILIGHTGHWGDRS 109
N.vitripennis         MFQLMDEIKAGLRYAFQTKNNLTLAISASGHGGLEAAIGNVLERGENMLIVKAGIWAERA 107
Apis_mellifera        FCMLMDEIKEGLQYIFQTNNRLTLALSASGHGGMEACLTNLLEPGETVLIVKSGIWSERA 116
C.elegans             FVQIMADVRLGLQYVFKTDNKYTFAVSGTGHSGMECAMVNLLEPGDKFLVVEIGLWGQRA 129
S.purpuratus          TLKIMDDIKKGLQYVFQTKNELTFAVSGSGHCGMEAAMMNLIEPGDVILVASNGIWGERI 113
Cyanothece_bacter     YLEMMDQIQDLLRYVWQTENELTISISGTGSAGMEASLANVVEPGDVVLIGVMGYFGHRL 104
Pr.coccus_marinus     YVDLMSEVQELLRYAWQTSNRLTLPMSGTGSAAMEATLANVVEPEDTVLVAIKGYFGHRL 140
                         :* :::  ::* ::* *  :: ::... .:*  : * :*  : .:    * :. *
                                                                          R
H.sapiens             VDIGERIGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLLFLTHGESSTGVLQP-LDGFGE 171
M.mulatta             MDIGERMGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLLFLTHGESSTGVLQP-LDGFGE 171
Pongo_pygmaeus        VDIGERMGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLLFLTHGESSTGVLQP-LDGFGE 171
C.jacchus             ADIGERLGARVHPMTKDPGGHYTLQEVEEGLAQHKPVLLFLTHGESSSGVLQP-LDGFGE 193
Canis_familiaris      ADIGERIGARVHSMVKDPGDYYTLQDVEEGLAQHKPALLFLTQGESSCGVLQP-LDGYGD 193
Felis_catus           ADIGERIGARVHPMIKDPGNHYTLQELEEALAQHKPVLLFLTQGESSSGVLQP-LDGYGE 193
Equus_caballus        ADIGERIGARVHMMVKDPGGHHTLREVEEGLAQHKPVLLFLTQGESSSGVLQP-LNGYGE 239
Bos_taurus            QDIAERIGARVYPMVKAPGGHFTLQEVEEALARHKPALLFLAHGESSTGVLQP-LDGYGE 193
O.cuniculus           AEVGERIGARVHPMIKDPGSHYTLQEVEECLAQHKPVLLFLTHGESSTGVLQP-LDGFGE 171
O.anatinus            ADIAERIGGKVHQLVKSPGGYYTLQDIEKGLIQHKPVLLFLTQGESSSGVLQP-LDGYGD 171
Rattus_norvegicus     AEIAERIGARVHQMIKKPGEHYTLQEVEEGLAQHKPVLLFLTHGESSTGVLQP-LDGFGE 171
Mus_musculus          AEIADRIGARVHQMIKKPGEHYTLQEVEEGLAQHKPVLLFLVHGESSTGVVQP-LDGFGE 193
Xenopus_laevis        GDIAERIGADVRYVSKPVGEAFTLKDVEKALAEHKPSLFFITHGESSSGVVQP-LDGLGD 194
Danio_rerio           SEIAERIGAKVNTVETMAGGYLTNEEIEKALNKYRPAVFFLTHGESSTGVVHP-IDGIGP 169
D.melanogaster        GDMARRYGAEVHYVEASFGRALSHEEITFAFEAHRPKVFFIAQGDSSTGIIQQNIRELGE 167
Aedes_aegypti         ADMATRYGADVRVVKSKVGQSLSLDEIRDALLIHKPSVLFLTQGDSSTGVLQG-LEGVGA 168
N.vitripennis         ADMAGRLGIRVDFMETEFGVAFGLREFELAVSEYRPKAVFVVHSESSAGLKQP-LEGLGD 166
Apis_mellifera        ACMANRLGANVKLLETDYTKGITLKELEIALEQHRPVVVFMVHAESSTGVKQP-LEDFGV 175
C.elegans             ADLANRMGIEVKKITAPQGQAVPVEDIRKAIADYKPNLVFVCQGDSSTGVAQP-LETIGD 188
S.purpuratus          ADLGKRLGANVKVLQNPTGVAFSLKQIEQSIHVYKPALFAITHGESSSGVLQP-LQGIGN 172
```

```
Cyanothece_bacter      VDMATRYGADVKTINKPWGENFSLAELKEAIETHKPAILGLVNAETSTGVRQP-LEGVGE 163
Pr.coccus_marinus      ADMAGRYKANVETIHKEWGNAFSLQEIEDALKKHTPAVLAIVHAETSTGVCQP-MDGIGD 199
                         :. *    *  :           :.   .  : *  . : :.::* *: :  :   *

H.sapiens              LCHRYKCLLLVDSVASLGGTPLYMDRQGIDILYSGSQKALNAPPGTSLISFSDKAKKKMY 231
M.mulatta              LCHRYKCLLLVDSVVSLGGTPLYMDQQGIDILYSGSQKVLNAPPGTSFISFSDKAKKKMY 231
Pongo_pygmaeus         LCHRYKCLLLVDSVVSVGGTPVYMDQQGIDILYSGSQKALNAPPGTSLISFSDKAKKKMY 231
C.jacchus              LCHRYKCLLLVDSVASLGGAPLYMDQQGIDILYSGSQKVLNAPPGTSLLSFSDTAKNKIY 253
Canis_familiaris       LCHRYKCLLLVDTVAALGGVPIYMDQQGIDVLYSGSQKVLNSPPGTSLISFSDKAKSKIY 253
Felis_catus            LCHRYNCLLLVDSVASLCGTPIYMDQQGIDVLYSGSQKVLNSPPGTSLISFSDKAKNKIY 253
Equus_caballus         LCHRYKCLLLVDSVASLGAVPIYMDQQGIDVLYSGSQKVLNAPPGTSLLSFSDKAKNKIY 299
Bos_taurus             LCHRHQCLLLVDSVASLGGAPVYMDQQGIDVLYSGSQKVLNAPPGTSLISFSNKAKNKIY 253
O.cuniculus            LCHRYKCLLLVDSVASLGGAPIYMDQQGIDVLYSGSQKALNAPPGTSLISFSDKAKSKIY 231
O.anatinus             LCHRYNCLLLVDCVASLGGTPIHMDKQAIDILYSGSQKVLNAPPGTSIISFSEKARKKIF 231
Rattus_norvegicus      LCHRYQCLLLVDSVASLGGVPIYMDQQGIDILYSGSQKVLNAPPGISLISFNDKAKSKVY 231
Mus_musculus           LCHRYQCLLLVDSVASLGGVPIYMDQQGIDIMYSSSQKVLNAPPGISLISFNDKAKYKVY 253
Xenopus_laevis         LCHRYNCLLLVDSVASLGGAPIYMDKQGIDILYSGSQKVLNAPPGTAPISFSEAASKKMF 254
Danio_rerio            LCRKYSCLFLVDSVAALGGAPICMDEQGIDILYTGSQKVLNAPPGTAPISFSERACHKIF 229
D.melanogaster         LCRKYDCFLIVDTVASLGGTEFLMDEWKVDVAYTGSQKSLGGPAGLTPISFSKRALTRIR 227
Aedes_aegypti          LCHQHNCLLIVDTVASLGGAPMFMDRWEIDAMYTGSQKVLGAPPGITPVSFSHRAVERYK 228
N.vitripennis          IVHKYGGLLIVDTVASLGAEPFFADAWGIDVVYTGSQKALGAPAGITPISFSPAAEKKIL 226
Apis_mellifera         LIHKYNALLIVDVVASLCGEPFFMDSWDIDAAYAGSQKAIGAPPGLAPISFSPRAEKKLF 235
C.elegans              ACREHGALFLVDTVASLGGTPFAADDLKVDCVYSATQKVLNAPPGLAPISFSDRAMEKIR 248
S.purpuratus           VCRRNDCLFLVDSVASLGGVPMFVDEWGIDAIYSGSQKVLGAPPGTAPISFSQRAISKFK 232
Cyanothece_bacter      LCREHNCLLLVDAVTSLGGVPFYTDKWGVDLAYSCSQKGLGCPPGLSPFTMSDRAIEKLQ 223
Pr.coccus_marinus      LCRKYNCLLLVDTVTSLGGVPLYLDEWKIDLAYSCSQKGLSCPPGLGPFSMNERAENKMS 259
                         :.   :::** *.:: .  .  *   :*  *: :** :. *.*   .::.  *  :

H.sapiens              SRKTKPFSFYLDIKWLANFWGCDD-QP-RMYHHTIPVISLYSLRESLALIAEQGLENSWR 289
M.mulatta              SRKTKPFSFYLDIKWLANFWGCDG-QP-RMYHHTIPVISLYSLRESLALIAEQGLENSWR 289
Pongo_pygmaeus         SRKTKPFSFYLDIKWLANFWGCDD-QP-RMYHHTIPIISLYSLRESLALIAEQGLENSWR 289
C.jacchus              SRKTKPSSFYLDVKYLANLWGCDG-QP-RMYHHTTPVVSLYSLREGLALLSEQGLENSWR 311
Canis_familiaris       ARKTKPVSFYLDMKWLANIWGCDD-QP-RVYHHTTPVISLYSLRESLALIAEQGTDARAR 311
Felis_catus            TRKTKPVSFYLDMKWLANIWGCDG-KP-RIYHHTTPVVSLYSLRESLALIAEQGLENSWR 311
Equus_caballus         TRKTKPFSFYLDIKWLANFWGCDD-QP-RMYHHTTPITGLYSLRESLALIAEQGLEDSWR 357
Bos_taurus             ARKTKPVSFYLDIQWLANFWGCDD-KP-RTYHHTTPVTSLYSLRESLAHLAEQGLENSWR 311
O.cuniculus            ARKTKPFSFYMDVQLLANIWGCDG-KP-RMYHHTTPVIGIFALRESLALLVEQGLEKSWQ 289
O.anatinus             SRKTKPVSFALDMSWLANFWGCDD-KP-RIYHHTTPVISLYSLREGLAILAEQGLEKSWK 289
Rattus_norvegicus      SRKTKPVSFYTDITYLSKLWGCEG-KT-RVIHHTLPVISLYCLRESLALISEQGLENSWR 289
Mus_musculus           SRKTKPVSFYTDITYLAKLWGCEG-ET-RVIHHTTPVTSLYCLRESLALIAEQGLENCWR 311
Xenopus_laevis         GRKTKPPSLYVDINWLANYWGCDG-KP-RIYHHTGPVTNFFTLREGLAILAELGLERSWA 312
Danio_rerio            NRKTKPISYFLDLNWLANYWGCDD-KPVRSYHHTGPVSSFYALRESLAILAETGLENSWK 288
D.melanogaster         KRKTKPKVYYFDILLIGQYWGCYG--TPRIYHHTISSTLLYGLREALAHFCAVGLKAVVR 285
Aedes_aegypti          RRNTKVKVYYWDMSLVGDYWGCFG--RPRIYHHTISSTLLYGLREAIAMACEEGLPALIA 286
N.vitripennis          TRKSPVPVFYWDMTWLGRYWNCFDPRTPRPYHHTISATLVYGLREALAQLAEEGLAASWA 286
Apis_mellifera         ERKTKPSSFYWDLTILGNYWKCFG-NEDRVYHHTMSATLLYGLREALAEIAEEGLQALWN 294
C.elegans              NRKQRVASFYFDAIELGNYWGCDG--ELKRYHHTAPISTVYALRAALSAIAKEGIDESIQ 306
S.purpuratus           SKKIRVPSFYLDLNHLANYWGCDA-GP-RRYHHTCPVTNLYQLREGLAMVAEEGLEECWK 290
Cyanothece_bacter      KRTTSVSNWYLDMNLLSQYWGEPR-----KYHHTAPCNMNYGLREALALIVEEGLENCWQ 278
Pr.coccus_marinus      NRKDKVPNWYLDVSLLNKYWGSDR-----VYHHTAPVNMNFGIREALRLLAEEGLEVSWG 314
                        :.        *   :   *           *** .    : :* .:      *

H.sapiens              QHREAAAYLHGRLQALGLQLFVKDPALRLPTVTTVAVPAGYDWRDIVSYVIDHFDIEIMG 349
M.mulatta              QHRETTAYLHGRLQALGLQLFVKDPALRLPMVTTVVVPAGYDWRDIVSYVLDHFDIEIMG 349
Pongo_pygmaeus         KHREAAAYLHGRLQALGLQLFVKDPALRLPTVTTVAVPAGYDWRDIVSYVMDHFDIEIMG 349
C.jacchus              KHREAAAYLHGRLQALGLRLFVKDPALRLPTVTTVAVPTGYDWRDIVSYLIERFGIEITG 371
Canis_familiaris       GAGGRTPRAP---SSAGPEPSRVPQAARLPTVTTVAAPAGYDWRDIVNYVMDHFDIEITG 368
Felis_catus            QHREVTAYLHGRLQGLGLQLFVKDPALRLPTVTTVAVPAGYDWRDIVNYVMDHFDIEITG 371
Equus_caballus         QHRETMAYLHGRLQGLGLRLFVKDPALRLPTITTVAVPAGYNWRDITNYVMDHFDIEITG 417
Bos_taurus             QHREASEYLHACLQGLGLQLFVKDPALRLPTITSVVVPTGYDWRDIVKYIMDHHDIEIAG 371
O.cuniculus            RHREVAQHLYRRLQELGLQLFVKDPALRLPTVTTVIVPASYRWRDIVSYVMHHFGIEITG 349
O.anatinus             HHEEVTQYLYEELQKLGLKLFVKEPAVRLPTVTTVAVPEGYNWKDIVDFLMNKYGIEITG 349
Rattus_norvegicus      RHREATAHLHKCLRELGLKFFVKDPEIRLPTITTVTVPAGYNWRDIVSYVLDHFNIEISG 349
Mus_musculus           RHREATAHLHKHLQEMGLKFFVKDPEIRLPTITTVTVPAGYNWRDIVSYVLDHFSIEISG 371
Xenopus_laevis         VHQENALKLHKGLEALGIKLFVKDPALRLPTVTTISVPNGYEWKDITTFIMKNHAIEITG 372
Danio_rerio            RHKEVAEYFHKGLEQMGLKLFVQDKKARLPTVTTIVAPPGYDWREITGYIMKTFNIEISG 348
D.melanogaster         RHQECSKRLQLGIEELGLEMFVSREEERLPTVNTIKVPFGVDWKKVAEYAMRKYSVEISG 345
Aedes_aegypti          RHEDCAKRLYRGLQDAGFELYAD-PKDRLSTVTTIKVPQGVDWLKAAQYAMKTYLVEISG 345
N.vitripennis          RHANVSQKFHEGLARRGYQLFVKQPQHRLKTVTAIMLPDGVEAQPIIRYAMDRYNLEFSG 346
Apis_mellifera         RHAAAAVRLRKGLELRGLQSYVKIPQYQLSTVISVQLPPGVVDKVVIQRAMEKYKVEISR 354
C.elegans              RHKDNAQVLYATLKKHGLEPFVVDEKLRLPCLTTVKVPEGVDWKDVAGKMMTN-GTEIAG 365
S.purpuratus           RHKAAARALYAGLEKLGLKLFVEDEATRLPTVSAILVPPGTDWKKVVVYVMQKYRIEISG 350
Cyanothece_bacter      RHQQNAELLWEGLEKLGLVCHVE-KQYRLPTLTTVRIPEGVNGKAVSIYLLKEYNIEIGG 337
```

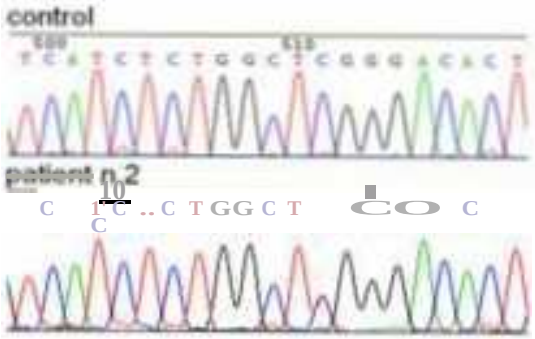
```
Pr.coccus_marinus      RHRTNAKSLWNSLENIGLELHVK-EELRLPTLTTVKIPEGLDGKAFTKHLLNNFGVEIGG 373
                                        *          :*  : ::  * .          :     *:

H.sapiens              GLGPSTGKVLRIGLLGCNATRENVDRVTEALRAALQHCPKKKL------ 392
M.mulatta              GLGPSAGKVLRIGLLGCNATRENVDRVTEALGAALQHCPKNKL------ 392
Pongo_pygmaeus         GLGPPTGKVLRIGLLGYNATHENVDRVTEALRAALQHCPKKKL------ 392
C.jacchus              GLGPSTGKVLRIGLMGCNATRENVDLVTEALREALQHCPKKKL------ 414
Canis_familiaris       GLGPSVGKVLRIGLLGCNATRENVDRVTHALQEALRHCPRHKL------ 411
Felis_catus            GLGPSMGKVLRIGLLGCNATRENVDRVIQALQEALQRCSRNKL------ 414
Equus_caballus         GLGPSMGKVLRIGLMGCNATRENVDRVIEALREALQRCPRNKL------ 460
Bos_taurus             GLGPSAGKVLRIGLLGGNATRESVDRVTRALREALQRCPRSKL------ 414
O.cuniculus            GLGPSADKVLRIGLLGCNATRENVDRLATALREALQHCAQSQL------ 392
O.anatinus             GLGPSIGMVLRIGLLGYNCSRANVDKVIQALREALKQCPKNKI------ 392
Rattus_norvegicus      GLGPSEDKVLRIGLLGYNATTENADRVAEALREALQHCPKNKL------ 392
Mus_musculus           GLGPTEERVLRIGLLGYNATTENVDRVAEALREALQHCPKNKL------ 414
Xenopus_laevis         GLGPSTGKVLRIGLMGYNSTQLNVDRVLEALRDALQQCPKKKM------ 415
Danio_rerio            GLGPSAGMVLRVGLMGCNSNKANVDKVLEALADALKHCHKSRV------ 391
D.melanogaster         GLGPTVEHVFRIGLMGENATVERVDMVLSILNEAIQSSKLGIKTDLSKI 394
Aedes_aegypti          GLGPTAGQVFRIGLMGQNATTERVDRVLQVFQEAVAAVKPDVQVKG-KM 393
N.vitripennis          GLGPTAGKVIRIGLMGVNATSGHADLVLRGLDDAVKYAKSCQLKSNI-- 393
Apis_mellifera         GLGPTVGKFLRIGLLRVNATPEKVDLVLRGLDEGLKYASESKL------ 397
C.elegans              GLGATVGKIWRIGTFGINSNSTKIENVVELLSKSIGEKSK--------- 405
S.purpuratus           GLGPSSGKVWRIGLMGYNANQDNVKRVLRVLEDGLNHSRTTSRL----- 394
Cyanothece_bacter      GLGELAGKVWRIGLMGYNSRPENVLLLLEALRKVL-------------- 372
Pr.coccus_marinus      GLGDLAGKVWRIGLMGYNSTSENVDKIINIFETELPKFR---------- 412
                       ***     . *:* :  *.       :   :   :
```



c.508G>A leading to p.Giy170Arg in patient 1



c.139G>A leading to p.Giy47Arg in patient 1



c.242C>T leading to p.Ser81Leu in patient 2